UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/510,209	10/05/2004	Yuzo Maruta	1163-0604PUS1	8655

2292 7590 03/05/2009
BIRCH STEWART KOLASCH & BIRCH
PO BOX 747
FALLS CHURCH, VA 22040-0747

EXAMINER	
ARMSTRONG, ANGELA A	

ART UNIT	PAPER NUMBER
2626	

NOTIFICATION DATE	DELIVERY MODE
03/05/2009	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

mailroom@bskb.com

Office Action Summary	Application No.	Applicant(s)	
	10/510,209	MARUTA, YUZO	
	Examiner	Art Unit	
	ANGELA A. ARMSTRONG	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *04 December 2008*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,2 and 4-7* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1,2 and 4-7* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

This Office Action is in response to the amendment filed December 4, 2008, amending claims 1, 2, 4, 6, and 7, and cancelling claim 3. Currently, claims 1-2 and 4-7 are pending.

Claim Rejections - 35 USC § 102

1. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

2. Claim 4 are rejected under 35 U.S.C. 102(b) as being anticipated by Kitaoka *et al.* (hereinafter "Kitaoka"), US Patent App. Pub. 2002/0010579.

Regarding claim 4, Kitaoka teaches a voice recognition index generator comprising: a representative word selector that selects single word as a representative word from an original set composed of a plurality of words and an acoustically similar word grouper that extracts from the original set, a word in which the acoustic likelihood between a sound feature vector for the word and a sound feature vector for the representative word is not less than a predetermined threshold, and including the extracted word in a same group as the representative word (Kitaoka teaches at paragraph 32, "the speech recognition apparatus generates and stores the similar sound group of the specific word beforehand…similar sound group includes reference patterns corresponding to sounds which are different from but similar to that of the specific word…rerecognition of the speech signal is performed by using the similar sound group of the specific word"; paragraph 30, teaches "pattern matching section performs pattern matching between each of the reference

patterns in a vocabulary stored in the dictionary section and the time-series data of the LPC cepstrum coefficients…similarity…likelihood ratio…between each of the reference patterns and each of the segments is computed."; paragraph 31 teaches, "pattern matching section selects as candidate words one or more words corresponding to the reference patterns which have high similarities with the LPC cepstrum coefficients"; and paragraph 45 teaches, "probability that the input speech signal actually represents the specific word…pattern matching section outputs a candidate word other than the specific word as the result of the recognition, if the received absolute level of confidence is equal to or lower than a predetermined reference level…reference level is experimentally determined beforehand"); and

an original-set replacer that passes to the representative word selector the word set left by removing from the original set the word affiliated by the group, as another original set to be processed by the representative word selector (paragraph 33 teaches "apparatus further generates reference patterns corresponding to sounds similar to that of a second specific word…second specific word is a word which means the opposite to the specific word…generated reference patterns are added to the similar sound group").

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 1, 5 and 7 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kitaoka *et al.* in view of Khan *et al.* (hereinafter "Khan"), US Patent App. Pub. 2002/0111810.

Regarding claim 1, Kitaoka teaches a voice recognition device for a car navigation system, comprising:

a sound analyzer that acoustically analyzes a user's vocal utterance inputted by a voice input means and for outputting a feature vector for the input sound (paragraph 28 teaches an acoustic analysis section, and paragraph 29 teaches a feature extraction section);

an acoustic-model storage that stores in advance respective acoustic models for predetermined sound units, either a syllable or a phoneme being deemed a sound unit (paragraph 30 teaches, "pattern matching between each of reference patterns in a vocabulary stored in the dictionary section and time-series data of the LPC cepstrum coefficients");

a sound-unit recognizer that checks the input-sound feature vector against the acoustic models to output a correlated sound-unit recognition candidate string (paragraphs 30-31, "the time-series data is divided into segments by using hidden Markov models and the similarity (i.e., likelihood ratio) between each of the reference patterns and each of the segments is computed...[e]ach of the reference patterns is a time-series of LPC cepstrum coefficients which are computed beforehand and correspond to one of words which should be identified");

Kitaoka does not explicitly teach, but Khan suggests, a word-and-position-information registration unit that correlates and registers in a word-and-position-information correlation dictionary the sound-unit recognition candidate string and position information acquired from a main unit of the car navigation system (Khan, Abstract, teaches a "navigation system includes an automatic speech recognition program that matches spoken words that describes geographic

features...to entries in a word list...geographic features closest to a certain position of a vehicle in which the navigation system is installed...[a]s the vehicle travels through a geographic area, the word list is rebuilt to include entries that correspond to the named geographic features closest to the new current vehicle position"; paragraph 51 teaches "name pronunciation data associated with those represented features that are closest to the current vehicle position").

Kitaoka in combination with Khan teaches, a position-information searcher/outputter that calculates acoustic likelihoods by collating the input-sound feature vector outputted by the sound analyzer, against sound feature vectors for the sound-unit recognition candidate strings in the word-and-position-information correlation dictionary, and outputting, to the car navigation main unit, position information associated with that sound-unit recognition candidate string whose calculated acoustic likelihood is not less than a predetermined threshold (Kitaoka teaches at paragraph 30, teaches "pattern matching section performs pattern matching between each of the reference patterns in a vocabulary stored in the dictionary section and the time-series data of the LPC cepstrum coefficients…similarity…likelihood ratio…between each of the reference patterns and each of the segments is computed"; paragraph 31 teaches, "pattern matching section selects as candidate words one or more words corresponding to the reference patterns which have high similarities with the LPC cepstrum coefficients"; and paragraph 45 teaches, "probability that the input speech signal actually represents the specific word…pattern matching section outputs a candidate word other than the specific word as the result of the recognition, if the received absolute level of confidence is equal to or lower than a predetermined reference level…reference level is experimentally determined beforehand"; Khan teaches word-and-position information and at paragraph 53, teaches "active word list that includes entries for named geographic features

that are close to the vehicle position...active word list...have a plurality of entries...[e]ach entry represents the phonetic pronunciation of the name of a particular represented geographic feature"; paragraph 58 teaches, "the geographic database is organized in a manner that facilitates finding the name pronunciation data for geographic features spatially...facilitate identifying name pronunciation data for geographic locations based upon the proximity of the geographic data from a selectable position"; paragraph 69, "name pronunciation data in the active word list...available for use by the automatic speech recognition program...threshold monitor…obtaining a new vehicle position...active word list").

It would have been obvious for one of ordinary skill in the art to combine the teaching elements of Kitaoka and Khan to include word-and-position information because Khan teaches his method has several advantages including "improved performance (as measured by reduced processing time and reduced memory requirements) of ASR algorithms operating in an in-vehicle environment" (paragraph 84).

Regarding claim 5, Kitaoka does not, but Khan suggests wherein the position-information searcher/outputter includes a voice recognition index-searching device, and uses the voice recognition index-searching device to search for and output words, their pronunciations, and position information stored in the word-and-position-information correlation dictionary or an external storage device (paragraph 53, teaches “active word list that includes entries for named geographic features that are close to the vehicle position...active word list...have a plurality of entries...[e]ach entry represents the phonetic pronunciation of the name of a particular represented geographic feature"; paragraph 58 teaches, "the geographic database is organized in

a manner that facilitates finding the name pronunciation data for geographic features spatially...facilitate identifying name pronunciation data for geographic locations based upon the proximity of the geographic data from a selectable position"; paragraph 69, "name pronunciation data in the active word list...available for use by the automatic speech recognition program...threshold monitor…obtaining a new vehicle position...active word list").

It would have been obvious for one of ordinary skill in the art to combine the teaching elements of Kitaoka and Khan to include word-and-position information because Khan teaches his method has several advantages including "improved performance (as measured by reduced processing time and reduced memory requirements) of ASR algorithms operating in an in-vehicle environment" (paragraph 84).

Regarding claim 7, Kitaoka teaches a car navigation system comprising:

a current position detector (paragraph 20, position detection unit);

a map data storage (paragraph 21, map data input unit);

an image display (paragraph 23, display unit);

a graphical pointer (paragraph 22 teaches, "control switches…mechanical switches… remote-control terminal"; paragraph 23 teaches "pointers which indicate the present position or traveling direction of the vehicle)"; and

a destination input device (paragraph 22, control switches).

The rest of the limitations of claim 7 are the same as or similar to those of claim 1, rejected above, and thus are rejected for the same reasons.

5. Claims 2 and 6 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kitaoka *et al.* in view of Khan *et al.*, and further in view of Ittycheriah *et al.* (hereinafter "Ittycheriah"), US Patent 6,192,337.

Regarding claim 2, Kitaoka and Khan do not explicitly teach, but Ittycheriah teaches: a confused-sound-unit matrix storage that stores in advance respective probabilities that an actual sound unit uttered by a human being will be recognized as a different recognition result as a consequence of the recognition precision of the sound analysis means, for each of recognition-result sound units (col. 8, ll. 49-67, teaches "distance measures calculated by the rejection processor for the comparisons between the newly uttered word and the existing words are preferably tabulated…tabular format may be organized in ranks based on an acoustic confusability threshold value...threshold value is set...any new word which results in a distance measure or score falling at or below the threshold value results in the newly uttered word being identified as likely to cause confusion with the associated existing word"); and

a word developer that outputs a candidate resembling the sound-unit recognition candidate string by replacing each sound unit in the sound-unit recognition candidate string outputted by the sound-unit recognition, with a recognition-result sound unit in which the probability that the confused-sound-unit matrix storage has stored for that sound unit is not less than a predetermined threshold (col. 8, ll. 49-67, "if the newly uttered word results in a distance measure falling above the threshold value, then the new word is identified as not likely to cause confusion with the associated existing word; col. 7, ll. 38-51, teaches "labeler outputs the symbols which comprise the predicted baseform...a leaf sequence corresponding to the predicted baseform is formed for the word uttered by the user"; col. 6, ll. 53-56, teaches "baseform and leaf

sequences...baseform of a word is a sequence of phonetic units (e.g., phones) that make up the word").

It would have been obvious for one of ordinary skill in the art at the time the invention was made to combine the teaching elements of Kitaoka and Khan with Ittycheriah to include a confused-sound-matrix because Ittycheriah teaches large vocabulary poses a problem to a user when a word is too similar to another one such that the speech recognizer is much less accurate on these words, if they appeared on the same list; a confused-sound-matrix would assist in handling this problem.

Kitaoka does not, but Khan suggests wherein the word-and-position-information registration correlates the resembling candidate to the position information acquired from the car navigation system main unit and registers this information in the word-and-position-information correlation dictionary (Khan teaches word-and-position information and at paragraph 48, teaches "threshold monitor routine obtains data indicating the current vehicle position…data indicating the current vehicle position may include the geographic coordinates of the vehicle position or alternatively, the data indicating the current vehicle position may be referenced to the map data contained in the geographic database that represent the road network"; paragraph 42 teaches "automatic speech recognition program matches the data representation of spoken words to one or more entries in an active word list (or dictionary)…performing…matching").

It would have been obvious for one of ordinary skill in the art to combine the teaching elements of Kitaoka and Khan to include word-and-position information because Khan teaches his method has several advantages including "improved performance (as measured by reduced

processing time and reduced memory requirements) of ASR algorithms operating in an in-vehicle environment" (paragraph 84).

Regarding claim 6, Kitaoka and Khan do not explicitly teach, but Ittycheriah suggests, wherein a word developer developing means extracts a probability stored in a confused-sound-unit matrix storage for each sound unit of the resembling candidate, and outputs a probability list for the resembling candidate (col. 7, line 62 – col. 8, line 9, teaches "comparing the newly uttered word to all existing vocabulary words to determine potential acoustic confusability…calculating respective distance measure or scores there between").

Kitaoka in combination with Khan and Ittycheriah suggests wherein the word-and-position-information registration unit correlates and registers in the word-and-position-information correlation dictionary both the probability list and the similar candidate with the position information (Kitaoka teaches at paragraph 32, "the speech recognition apparatus generates and stores the similar sound group of the specific word beforehand…similar sound group includes reference patterns corresponding to sounds which are different from but similar to that of the specific word…rerecognition of the speech signal is performed by using the similar sound group of the specific word"; Khan teaches word-and-position information and at paragraph 53, 58 and 69, as discussed above); and

wherein the position-information searcher/outputter, after reading a resembling word candidate stored in the word-and-position-information correlation dictionary and the probability list for that resembling word, and if the probability in its probability list is not less than a predetermined threshold, calculates the acoustic likelihood by checking the input-sound feature

vector against the sound feature vector outputted by a sound feature vector generator and outputs the sound-unit recognition candidate string whose acoustic likelihood is not less than the predetermined threshold, and if the probability in the probability list is less than the predetermined threshold, the position-information searcher/outputter uses the voice recognition index-searching device to search for words, their pronunciations, and position information stored in the external storage device (Kitaoka teaches at paragraph 30, teaches "pattern matching section performs pattern matching between each of the reference patterns in a vocabulary stored in the dictionary section and the time-series data of the LPC cepstrum coefficients...similarity...likelihood ratio...between each of the reference patterns and each of the segments is computed."; paragraph 31 teaches, "pattern matching section selects as candidate words one or more words corresponding to the reference patterns which have high similarities with the LPC cepstrum coefficients"; and paragraph 45 teaches, "probability that the input speech signal actually represents the specific word...pattern matching section outputs a candidate word other than the specific word as the result of the recognition, if the received absolute level of confidence is equal to or lower than a predetermined reference level...reference level is experimentally determined beforehand"; Khan teaches word-and-position information and at paragraph 53, 58 and 69, as discussed above).

It would have been obvious for one of ordinary skill in the art to combine the teaching elements of Kitaoka and Khan to include word-and-position information because Khan teaches his method has several advantages including "improved performance (as measured by reduced processing time and reduced memory requirements) of ASR algorithms operating in an in-vehicle environment" (paragraph 84).

Response to Arguments

Applicant's arguments filed December 4, 2008, have been fully considered but they are not persuasive.

Regarding claim 4, Applicant argues Kitaoka fails to disclose an original-set replacer as claimed. The Examiner cannot concur. Kitaoka (at paragraph 33) teaches the apparatus generates reference patterns corresponding to sounds similar to that of a second specific word…second specific word is a word which means the opposite to the specific word…generated reference patterns are added to the similar sound group and additionally teaches (paragraph 34) as an example, the word `NO` is selected as the second specific word and reference patterns corresponding to sounds similar to that of the word `NO` is also generated and added to the similar sound group. The reference patterns corresponding to sounds /au/, /uu/, and the like are added to the similar sound group in this case and thus it is preferable that the similar sound group should include the reference patterns corresponding to sounds similar to that of the second specific word. The system generates new reference patterns for the second specific word, which creates a new different database/word list/vocabulary/ than what was available originally.

Applicant argues Kahn fails to disclose that the spoken word or a candidate string matched from the spoken word is correlated and registered with position data as claimed. In response, the Examiner argues Kahn teaches the rebuilder routine builds the active word list with current position information and the name pronunciation data associated with the represented geographic features of the current vehicle position [paragraph 0050].

Conclusion

1. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to ANGELA A. ARMSTRONG whose telephone number is (571)272-7598. The examiner can normally be reached on Monday-Thursday 11:30-8:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick N. Edouard can be reached on 571-272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Angela A Armstrong/
Primary Examiner, Art Unit 2626